

CAMBRIDGE

国外电子信息精品著作(影印版)

高频CMOS模拟 集成电路基础

**Fundamentals of High-Frequency
CMOS Analog Integrated Circuits**

(China Edition)

**Duran Leblebici
Yusuf Leblebici**



科学出版社

国外电子信息精品著作（影印版）

**Fundamentals of High-Frequency CMOS
Analog Integrated Circuits**

高频 CMOS 模拟集成电路基础

Duran Leblebici

Yusuf Leblebici

科学出版社

北京

图字：01-2011-3631

内 容 简 介

本书以设计为核心理念从基础模拟电路讲述到射频集成电路的研发。系统地介绍了高频集成电路体系的构建与运行，重点讲解了晶体管级电路的工作体系，设备性能影响及伴随响应，以及时域和频域上的输入输出特性。本书适合电子信息专业的高年级本科生及研究生作为 RFCMOS 电路设计相关课程的教材使用，也适合模拟电路及射频电路工程师作为参考使用。

Fundamentals of High-Frequency CMOS Analog Integrated Circuits, first edition (0-521-51340-1) by Duran Leblebici, Yusef Leblebici first published by Cambridge University Press 2009

All rights reserved.

This reprint Chinese edition for the People's Republic of China is published by arrangement with the Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.

© Cambridge University Press & Science Press Ltd. 2011

This book is in copyright. No reproduction of any part may take place without the written permission of Cambridge University Press and Science Press Ltd.

This edition is for sale in the People's Republic of China (excluding Hong Kong SAR, Macau SAR and Taiwan Province) only.

此版本仅限在中华人民共和国境内（不包括香港、澳门特别行政区及台湾地区）销售。

图书在版编目(CIP)数据

高频 CMOS 模拟集成电路基础 = Fundamentals of High-Frequency CMOS Analog Integrated Circuits: 英文/(土)莱布莱比吉 (Leblebici, D.) 等著. —影印版. —北京: 科学出版社, 2011.7

(国外电子信息精品著作)

ISBN 978-7-03-031519-9

I. ①高… II. ①莱… III. ①CMOS 电路: 模拟集成电路-英文
IV. ①TN432

中国版本图书馆 CIP 数据核字 (2011) 第 112932 号

责任编辑: 孙伯元/责任印制: 赵 博/封面设计: 陈 敬

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2011 年 6 月第 一 版 开本: B5(720×1000)

2011 年 6 月第一次印刷 印张: 19 3/4

印数: 1—3 000 字数: 400 000

定价: 60.00 元

(如有印装质量问题, 我社负责调换)

《国外电子信息精品著作》序

20世纪90年代以来,信息科学技术成为世界经济的中坚力量。随着经济全球化的进一步发展,以微电子、计算机、通信和网络技术为代表的信息技术,成为人类社会进步过程中发展最快、渗透性最强、应用面最广的关键技术。信息技术的发展带动了微电子、计算机、通信、网络、超导等产业的发展,促进了生命科学、新材料、能源、航空航天等高新技术产业的成长。信息产业的发展水平不仅是社会物质生产、文化进步的基本要素和必备条件,也是衡量一个国家的综合国力、国际竞争力和发展水平的重要标志。在中国,信息产业在国民经济发展中占有举足轻重的地位,成为国民经济重要支柱产业。然而,中国的信息科学支持技术发展的力度不够,信息技术还处于比较落后的水平,因此,快速发展信息科学技术成为我国迫在眉睫的大事。

要使我国的信息技术更好地发展起来,需要科学工作者和工程技术人员付出艰辛的努力。此外,我们要从客观上为科学工作者和工程技术人员创造更有利于发展的环境,加强对信息技术的支持与投资力度,其中也包括与信息技术相关的图书出版工作。

从出版的角度考虑,除了较好较快地出版具有自主知识产权的成果外,引进国外的优秀出版物是大有裨益的。洋为中用,将国外的优秀著作引进到国内,促进最新的科技成就迅速转化为我们自己的智力成果,无疑是值得高度重视的。科学出版社引进一批国外知名出版社的优秀著作,使我国从事信息技术的广大科学工作者和工程技术人员能以较低的价格购买,对于推动我国信息技术领域的科研与教学是十分有益的事。

此次科学出版社在广泛征求专家意见的基础上,经过反复论证、仔细遴选,共引进了接近30本外版书,大体上可以分为两类,第一类是基础理论著作,第二类是工程应用方面的著作。所有的著作都涉及信息领域的最新成果,大多数是2005年后出版的,力求“层次高、内

容新、参考性强”。在内容和形式上都体现了科学出版社一贯奉行的严谨作风。

当然，这批书只能涵盖信息科学技术的一部分，所以这项工作还应该继续下去。对于一些读者面较广、观点新颖、国内缺乏的好书还应该翻译成中文出版，这有利于知识更好更快地传播。同时，我也希望广大读者提出好的建议，以改进和完善丛书的出版工作。

总之，我对科学出版社引进外版书这一举措表示热烈的支持，并盼望这一工作取得更大的成绩。



中国科学院院士

中国工程院院士

2006年12月

To
Yıldız Lelebici – wife, mother, and colleague.

Preface

In the first half of the twentieth century the radio was the main activity area of the electronics industry and, correspondingly, RF circuits occupied a considerable part in the electronic engineering curriculum and books published in this period. Properties of resonance circuits and electronic circuits using them, single-tuned amplifiers as the input stages of receivers, double-tuned circuits as IF amplifiers, RF sinusoidal oscillators and high-power class-C amplifiers have been investigated in depth. It must be kept in mind that the upper limit of the radio frequencies of those days was several tens of MHz, the inductors used in tuned circuits were air-core or ferrite-core coils with inductance values in the micro-henries to milli-henries range, having considerably high quality factors, ranging from 100 to 1000, and the tuning capacitors were practically lossless.

The knowledge developed for the vacuum-tube circuits easily adapted to the transistors with some modifications related to the differences of the input and output resistances of the devices. In the meantime, the upper limit of the frequency increased to about 100 MHz for FM radio and to hundreds of MHz for UHF-TV. The values of the inductors used in these circuits correspondingly decreased to hundreds to tens of nano-henries. But these inductors were still wound, high-Q discrete components.

In the second half of the twentieth century, the emergence of integrated circuits drastically increased the reach of electronic engineering. Digital electronics on one side and analog electronics using the potentials of the operational amplifiers on the other side forced the curricula and the textbooks to skip certain old and “already known” subjects (among them resonance circuits and tuned amplifiers), to open room to these new subjects. The development of (inductorless) active filters that replaced the conventional passive LC filters extensively used in telecommunication systems even decreased the importance of inductors.

Rapid development of CMOS technology and the steady decrease of the dimensions of the devices according to “Moore’s Law” led to an increase in the complexity of ICs (from now on, the VLSI circuits) and the operating frequencies as well. In the digital realm this helped to improve the performances of digital computers and digital telecommunication systems. In the analog realm the operation frequencies of the circuits increased to the GHz region and correspondingly the inductance values decreased to below 20 nH, so that now it was possible to realize them as on-chip components. Hence the freedom from external bulky discrete inductors opened a new

horizon towards small and light-weight mobile systems; the mobile telephone, GPS systems, Bluetooth, etc.

But there is a problem related to this development; the quality factors of on-chip inductors are very low, usually around 10–100, and tuning capacitors are not lossless any more. Most of the earlier theory (and design practice) that was developed with very high-Q discrete components does not easily translate to such integrated high-frequency circuits and on-chip components with less than ideal characteristics. Instead of relying on the comprehensive theory and analytical design of high-frequency analog circuits that was in place many decades earlier, most of the new-generation designers were tempted to adopt rather ad-hoc design strategies not grounded in sound theory, and to explain away the inevitable inconsistencies as “secondary effects”. To make matters even more complicated, systematic treatment of subjects such as high-frequency circuit behavior and resonance/tuned circuits has been missing from the electrical engineering curricula for several decades, and analog designers entering the field of RF/high-frequency design had to re-learn these subjects.

One of the objectives of this book is to fill this gap; to introduce the fundamental aspects of high-frequency circuit operation, to systematically discuss the behavior of key components (in particular, submicron MOSFETs and on-chip passive components), summarize the behavior of the series and parallel resonance circuits in detail and investigate the effects of the losses of on-chip inductors and capacitors that are usually not taken into account in formulas derived for high-Q resonance circuits.

Since all these circuits are being developed and used usually in the GHz range, sometimes close to the physical limits of the devices, it becomes necessary to recapitulate the behavior of MOS transistors together with their important parasitics and the frequency related secondary effects.¹ In Chapter 1 the basic current–voltage relations of MOS transistors are derived, taking into account the parabolic (not linear) shape of the inversion charge profile that leads to a different approach to understand the channel shortening effect and the gate-source capacitance of the transistor. The velocity saturation effect and the behavior of a MOS transistor under a velocity saturation regime are also investigated as an important issue for small geometry devices. Although it is not used extensively for HF applications, the sub-threshold regime is also investigated in brief.

In addition to the intrinsic behavior of MOS transistors the parasitics – that are inevitable and have severe effects on the overall behavior especially at high frequencies – have been discussed, mostly in connection with the BSIM3 parameters. The properties, limits and parasitics of the passive on-chip components, namely resistors, fixed value and variable capacitors (varactors) and inductors are also summarized in this chapter from a realistic and design-oriented point of view.

The subject of Chapter 2 is the DC properties of the basic analog MOS circuits that will be investigated in the following chapters. The interactive use of analytic

¹ This part of the book must not be considered as an alternative to the existing sound and comprehensive models such as EKV and others, but rather as an attempt to explain the behavior of MOS transistors, based on the basic laws of electrostatics and circuit theory that all electronics students already know.

expressions – that provide interpretable knowledge about the basic behavior of the circuit – and SPICE simulations – that give the designer a possibility to “experiment”, to fine-tune and optimize the circuit, all secondary and parasitic effects included – are exemplified throughout the chapter. It is believed that to have the ability of using together the analytical expressions and the power of SPICE is a “must” for an analog designer.

In Chapter 3 the frequency-dependent behaviors of the basic circuits are given, not limited to the frequency characteristics of the gain but including the input and output impedances. The important properties of them, usually not dealt with in books, are investigated and their effects on the performance of wide-band circuits are underlined. The basics of the techniques used to enhance the gain; the additive approach (distributed amplifiers) and the cascading strategies to reach to the wide-band amplifiers (not only voltage amplifiers, but also current amplifiers, transadmittance amplifiers and transimpedance amplifiers) are systematically investigated.

In Chapter 4 first the resonance circuits are recapitulated with this approach and the behavioral differences of high-Q and low-Q resonance circuits are underlined. Afterwards, tuned amplifiers are systematically investigated taking into account the low-Q effects, not only for single-tuned amplifiers but also for double-tuned and staggered tuned amplifiers that are not covered in many new (and even older) books in detail. The LNA, that is one of the most important classes of tuned amplifiers, is also investigated in this chapter together with the noise behavior of MOS transistors, that is developed with a different approach.

LC sinusoidal oscillators are discussed in Chapter 5 with the negative resistance approach and the classical positive feedback approach as well, with emphasis on the effects of the low-Q components. The problems related to the frequency stability of LC oscillators are discussed and the phase noise in LC oscillators is investigated with a different approach.

The last chapter is devoted to a summary of the higher-level system view of HF analog circuits, especially in the context that virtually all such high-frequency circuits are eventually integrated with considerable digital circuitry for interface, post-processing and calibration purposes – and that such integration is increasingly done on the same silicon substrate. The traditional system-level view of high-frequency components and circuits is strongly influenced by conventional (all-analog) modulation and transmission systems modeling, which is based almost exclusively on the frequency domain. The behavior of all digital systems, on the other hand, is preferably described in the time domain. While the translation between these two domains is (in theory) quite straightforward, the designers must develop a sense of how some of their choices in the analog realm eventually influence the behavior of the digital part, and vice versa. Data converters (analog-to-digital and digital-to-analog converters) naturally play an important role in this translation between domains, and Chapter 6 attempts to summarize the key criteria and parameters that are used to describe system-level performance.

The target audience of this book includes advanced undergraduate and graduate-level students who choose analog/mixed-signal microelectronics as their area of

specialization, as well as practicing design engineers. The required background that is needed to follow the material is consistent with the typical physics, math and circuits background that is acquired by the third (junior) year of a regular electrical and computer engineering (ECE) curriculum.

Solved design examples are provided to guide the reader through the decision process that accompanies each design task, emphasizing key trade-offs and eventual approximations.

A number of individuals have contributed with their time and their efforts to the creation of this textbook. In particular, both authors would like to thank Mrs Yildiz Leblebici who read the entire manuscript, carefully checked the analytical derivations throughout all chapters, and provided valuable insight as an experienced electronics teacher. The authors also acknowledge the generous support of Mr Giovanni Chiappano from austriamicrosystems A.G.² for offering the use of transistor parameters in numerous examples.

The idea of this book was originally launched with the enthusiastic encouragement of Dr Philip Meyler of Cambridge University Press, who saw the need for a design-oriented text in this field and patiently followed through its early development. We are deeply grateful to Dr Julie Lancashire, our publisher, for her guidance, support and encouragement over the years, leading up to the final stages of production. The editorial staff of Cambridge University Press has been wonderfully supportive throughout this project. We would especially like to thank Ms Sarah Matthews and Mr Christopher Miller for their valuable assistance.

Last but not least, the authors would also like to thank all reviewers who read all or parts of the manuscript and provided very valuable comments.

² Device characteristics and technology parameters provided by austriamicrosystems A.G. will be labeled with the acronym "AMS" through the text.

Contents

	<i>Preface</i>	<i>page xi</i>
1	Components of analog CMOS ICs	1
	1.1 MOS transistors	1
	1.1.1 Current–voltage relations of MOS transistors	3
	1.1.1.1 The basic current–voltage relations without velocity saturation	4
	1.1.1.2 Current–voltage relations under velocity saturation	11
	1.1.1.3 The sub-threshold regime	15
	1.1.2 Determination of model parameters and related secondary effects	19
	1.1.2.1 Mobility	20
	1.1.2.2 Gate capacitance	20
	1.1.2.3 Threshold voltage	21
	1.1.2.4 Channel length modulation factor	23
	1.1.2.5 Gate length (L) and gate width (W)	24
	1.1.3 Parasitics of MOS transistors	25
	1.1.3.1 Parasitic capacitances	26
	1.1.3.2 The high-frequency figure of merit	30
	1.1.3.3 The parasitic resistances	31
	1.2 Passive on-chip components	36
	1.2.1 On-chip resistors	36
	1.2.2 On-chip capacitors	38
	1.2.2.1 Passive on-chip capacitors	38
	1.2.2.2 Varactors	40
	1.2.3 On-chip inductors	43
2	Basic MOS amplifiers: DC and low-frequency behavior	49
	2.1 Common source (grounded source) amplifier	49
	2.1.1 Biasing	53
	2.1.2 The small-signal equivalent circuit	54
	2.2 Active transistor loaded MOS amplifier (CMOS inverter as analog amplifier)	63
	2.3 Common-gate (grounded-gate) amplifier	68

2.4	Common-drain amplifier (source follower)	70
2.5	The “long tailed pair”	75
2.5.1	The large signal behavior of the long tailed pair	84
2.5.2	Common-mode feedback	88
3	High-frequency behavior of basic amplifiers	95
3.1	High-frequency behavior of a common-source amplifier	97
3.1.1	The R-C load case	99
3.2	The source follower amplifier at radio frequencies	103
3.3	The common-gate amplifier at high frequencies	110
3.4	The “cascode” amplifier	114
3.5	The CMOS inverter as a transimpedance amplifier	118
3.6	MOS transistor with source degeneration at high frequencies	126
3.7	High-frequency behavior of differential amplifiers	129
3.7.1	The R-C loaded long tailed pair	129
3.7.2	The fully differential, current-mirror loaded amplifier	132
3.7.3	Frequency response of a single-ended output long tailed pair	136
3.7.4	On the input and output admittances of the long tailed pair	141
3.8	Gain enhancement techniques for high-frequency amplifiers	143
3.8.1	“Additive” approach: distributed amplifiers	144
3.8.2	Cascading strategies for basic gain stages	146
3.8.3	An example: the “Cherry–Hooper” amplifier	148
4	Frequency-selective RF circuits	155
4.1	Resonance circuits	156
4.1.1	The parallel resonance circuit	156
4.1.1.1	The quality factor of a resonance circuit	160
4.1.1.2	The quality factor from a different point of view	163
4.1.1.3	The “ Q enhancement”	164
4.1.1.4	Bandwidth of a parallel resonance circuit	168
4.1.1.5	Currents of L and C branches of a parallel resonance circuit	169
4.1.2	The series resonance circuit	170
4.1.2.1	Component voltages in a series resonance circuit	172
4.2	Tuned amplifiers	172
4.2.1	The common-source tuned amplifier	173
4.2.2	The tuned cascode amplifier	179
4.3	Cascaded tuned stages and the staggered tuning	181
4.4	Amplifiers loaded with coupled resonance circuits	189
4.4.1	Magnetic coupling	189
4.4.2	Capacitive coupling	194
4.5	The gyrator: a valuable tool to realize high-value on-chip inductances	194
4.5.1	Parasitics of a non-ideal gyrator	197
4.5.2	Dynamic range of a gyrator-based inductor	201

4.6	The low-noise amplifier (LNA)	202
4.6.1	Input impedance matching	203
4.6.2	Basic circuits suitable for LNAs	207
4.6.3	Noise in amplifiers	210
4.6.3.1	Thermal noise of a resistor	212
4.6.3.2	Thermal noise of a MOS transistor	213
4.6.4	Noise in LNAs	224
4.6.5	The differential LNA	234
5	L-C oscillators	237
5.1	The negative resistance approach to L-C oscillators	237
5.2	The feedback approach to L-C oscillators	245
5.3	Frequency stability of L-C oscillators	249
5.3.1	Crystal oscillators	251
5.3.2	The phase-lock technique	253
5.3.3	Phase noise in oscillators	255
6	Analog–digital interface and system-level design considerations	259
6.1	General observations	259
6.2	Discrete-time sampling	263
6.3	Influence of sampling clock jitter	265
6.4	Quantization noise	267
6.5	Converter specifications	268
6.5.1	Static specifications	269
6.5.2	Frequency-domain dynamic specifications	273
6.6	Additional observations on noise in high-frequency ICs	275
	<i>Appendix A Mobility degradation due to the transversal field</i>	277
	<i>Appendix B Characteristic curves and parameters of AMS 0.35 micron NMOS and PMOS transistors</i>	279
	<i>Appendix C BSIM3-v3 parameters of AMS 0.35 micron NMOS and PMOS transistors</i>	281
	<i>Appendix D Current sources and current mirrors</i>	287
	D.1 DC current sources	287
	D.2 Frequency characteristics of basic current mirrors	289
	D.2.1 Frequency characteristics for normal saturation	291
	D.2.2 Frequency characteristics under velocity saturation	292
	References	293
	Index	297

1 Components of analog CMOS ICs

1.1 MOS transistors

The basic structure of an n-channel metal–oxide–semiconductor (NMOS) transistor built on a p-type substrate is shown in Fig. 1.1. The MOS transistor consists of two disjoint p-n junctions (source and drain), bridged by a MOS capacitor composed of the thin gate oxide and the poly-silicon gate electrode. If a positive voltage with sufficiently large magnitude is applied to the gate electrode, the resulting vertical electric field between the substrate and the gate attracts the negatively charged electrons to the surface. Once the electron concentration on the surface exceeds the majority hole concentration of the p-type substrate, the surface (the channel) is said to be inverted, i.e., a conducting channel is formed between the source and the drain. The carriers, i.e., the electrons in an NMOS transistor, enter the channel region underneath the gate through the source contact, leave the channel region through the drain contact, and their movement in the channel region is subject to the control of the gate voltage.

To ensure that both p-n junctions are continuously reverse biased, the substrate potential is kept lower than the source and drain terminal potentials. Note that the device structure is symmetrical with respect to the drain and source regions; the different roles of these two regions are defined only in conjunction with the applied terminal voltages and the direction of the drain current. In an n-channel MOS (NMOS) transistor, the source is defined as the n+ region which has a lower potential than the other n+ region, the drain. This means that the current flow direction is from the drain to the source. By convention, all terminal voltages of the device are defined with respect to the source potential.

The value of the gate-to-source voltage (V_{GS}) necessary to cause surface inversion (to create the conducting channel) is called the threshold voltage V_T . This quantity depends on various device and process parameters such as the work function difference between the gate and the substrate, the substrate (surface) Fermi potential, the depletion region charge concentration, the interface charge concentration, the gate oxide thickness and oxide (dielectric) permittivity, as well as the concentration of the channel implantation that is used to adjust the threshold voltage level.

If the applied gate-to-source voltage exceeds the threshold voltage of the MOS transistor, a sufficiently high concentration of electrons is achieved in the channel region, leading to surface inversion. Thus, an n-type conducting channel is formed between the source and the drain, which is capable of carrying the drain (channel) current. If a small positive voltage is applied to the drain, a current proportional to this voltage will start to flow from the drain to the source through the conducting channel. The effective

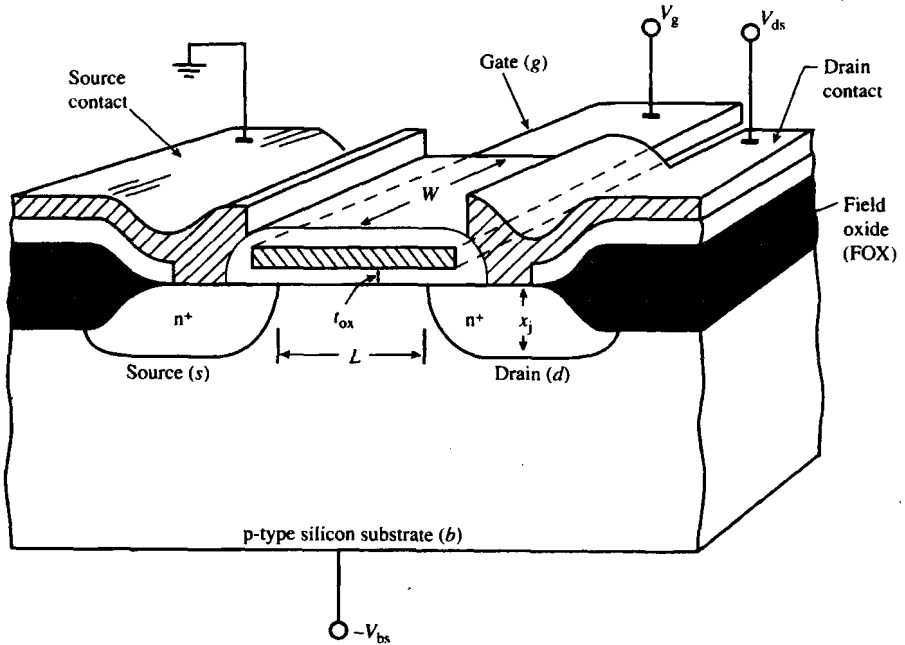


Figure 1.1 Simplified cross-section view of an n-channel MOS (NMOS) transistor (after Taur and Ning [13]).

resistivity of the continuous inversion layer between the source and the drain depends on the gate voltage. This operating mode is called the linear (or triode) mode, where the channel region acts as a voltage-controlled resistor. During this operating mode, the electron velocity in the channel is usually much lower than the drift velocity limit.

As the applied drain voltage is increased, the inversion layer charge and the channel depth at the drain end start to decrease. Eventually, when the drain voltage reaches a limit value called the saturation voltage ($V_{D(sat)}$), the inversion charge at the drain is reduced – theoretically – to zero, and the velocity of electrons – theoretically – reaches very high values, as discussed in the following sections. This event is named as the “pinch-off” of the channel. Beyond the pinch-off point, i.e., for drain voltage values larger than the saturation voltage, electrons travel in a very shallow pinched-off channel with a very high velocity, which is called the “saturation velocity”. This operating regime is known as the saturation mode.

If the transistor is formed on an n-type substrate, using two p+ regions as source and drain, this structure is called a p-channel MOS (PMOS) transistor. In a PMOS transistor, the fundamental mechanisms of surface inversion and channel conduction are exactly the same as in NMOS transistors, although the majority of carriers consist of holes, not electrons. Thus, the gate-to-source voltage applied to the gate electrode to achieve surface inversion must be negative. Also, it should be taken into account that the hole mobility is considerably smaller than the electron mobility at room

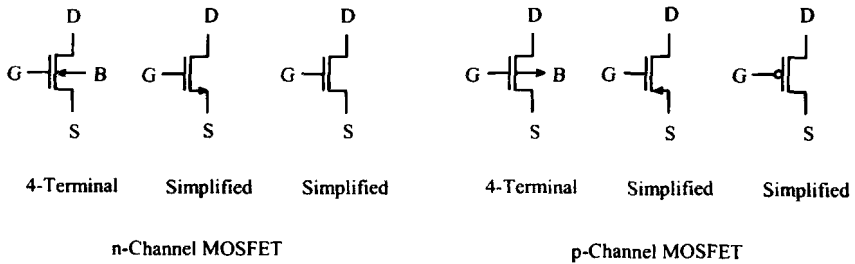


Figure 1.2 Commonly used circuit symbols for NMOS and PMOS transistors.

temperature, which leads to a smaller effective channel conductance for the PMOS transistor with the same channel dimensions. Nevertheless, the complementary nature of NMOS/PMOS biasing and operating conditions offers very useful circuit implementation possibilities, which underlines the importance and the wide-spread use of complementary MOS (CMOS) circuits in a very large range of applications.

Commonly used circuit symbols for n-channel and p-channel MOS transistors are shown in Fig. 1.2. While the four-terminal representation shows all external terminals of the device, the three-terminal symbol is usually preferred for simplicity. Unless noted otherwise, the substrate terminals are always assumed to be connected to the lowest potential for NMOS devices, and to the highest potential for PMOS devices.

1.1.1 Current–voltage relations of MOS transistors

The basic (so-called Level-1) current–voltage relations of a MOS transistor are given in most basic electronics textbooks. Since these relations contain a small number of parameters, they are convenient for hand calculations. The parameters of these expressions are:

- the mobility of electrons (or holes), μ ;
- the gate capacitance per unit area, C_{ox} ;
- the threshold voltage of the transistor, V_T ;
- the gate-length modulation coefficient, λ ;
- and the aspect ratio of the transistor, (W/L) .

In the following, these relations are derived with a different approach, to remind the reader of the fundamentals, and also to clarify the understanding of device behavior. In addition, the derivation presented here is based on a realistic profile of the channel-region inversion charge (as calculated from the fundamental electric field expressions), as opposed to the classical gradual channel approach which assumes linear charge profiles in the channel. This model represents the transistor under moderate to strong inversion conditions with reasonable accuracy for hand calculations, provided that the channel length is not too short and the transistor is not in the velocity saturation region.

For short-channel MOS transistors in which the carrier velocities reach saturation, i.e. approach a limit velocity, this model is no longer valid. Since a great majority of transistors realized in analog MOS integrated circuits today have channel lengths in

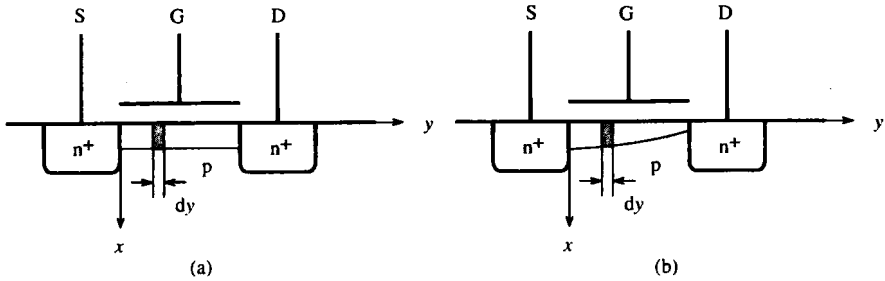


Figure 1.3 Inversion channel profiles of an NMOS transistor for (a) $V_{GS} > V_T$, $V_{DS} = 0$, (b) $V_{GS} > V_T$, $V_{DS} > 0$.

the sub-half-micron range, they may easily enter the velocity saturation region. Therefore it is necessary to derive rules to check whether a transistor is operating in the velocity saturation region or not, and to obtain expressions that are valid for velocity saturated transistors.

1.1.1.1 The basic current–voltage relations without velocity saturation

The cross-section of an NMOS transistor having an inversion layer along the channel owing to a gate–source voltage greater than the threshold voltage and zero source–drain voltage is shown in Fig. 1.3(a). Since there is no surface inversion for gate voltages smaller than the threshold voltage V_T , the value of the inversion charge density is

$$Q_i = -C_{ox}(V_{GS} - V_T) \text{ [coulomb/cm}^2\text{]}$$

where $(V_{GS} - V_T)$ is the “effective gate voltage” for this case. The amount of the inversion charge of a transistor having a channel length L and a channel width W is

$$\bar{Q}_i = -C_{ox}WL(V_{GS} - V_T)$$

The minus signs in front of these expressions denote that this is a negative charge, since the carriers in the inversion layer of an NMOS transistor are electrons.

When we apply a positive drain voltage with respect to the source, a drain current (I_D) flows in the $-y$ direction and is constant along the channel (Fig. 1.3(b)). However, owing to the voltage drop on the channel resistance, the voltage along the channel is not constant. Although the effective gate voltage – inducing the inversion charge – at the source end of the channel is $(V_{GS} - V_T)$, it decreases along the channel and becomes equal to $(V_{GD} - V_T) = (V_{GS} - V_{DS} - V_T)$ at the drain end. If the channel voltage with respect to the source is denoted by $V_c(y)$, the effective gate voltage as a function of y can be written as

$$V_{\text{eff}}(y) = (V_{GS} - V_T) - V_c(y) \quad (1.1)$$

and the amount of the inversion charge in an infinitesimal channel segment dy is

$$d\bar{Q}_i(y) = -C_{ox}W[(V_{GS} - V_T) - V_c(y)]dy \quad (1.2)$$